View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Am. J. Hum. Genet. 68:927–936, 2001

A Genomewide Screen in Multiplex Rheumatoid Arthritis Families Suggests Genetic Overlap with Other Autoimmune Diseases

Damini Jawaheer,[1] Michael F. Seldin,[2] Christopher I. Amos,[3] Wei V. Chen,[3] Russell Shigeta,[2] Joanita Monteiro,[1] Marlene Kern,[1] Lindsey A. Criswell,[4] Salvatore Albani,[5] J. Lee Nelson,[6] Daniel O. Clegg,[8] Richard Pope,[9] Harry W. Schroeder Jr.,[10] S. Louis Bridges Jr.,[10] David S. Pisetsky,[11] Ryk Ward,[12] Daniel L. Kastner,[13] Ronald L. Wilder,[13] Theodore Pincus,[15] Leigh F. Callahan,[16] Donald Flemming,[14] Mark H. Wener,[7] and Peter K. Gregersen[1,*]

[1]Division of Biology and Human Genetics, North Shore University Hospital, Manhasset, NY[†]; [2]Department of Biological Chemistry, University of California at Davis, Davis, CA[†]; [3]Departments of Epidemiology and Biomathematics, University of Texas M. D. Anderson Cancer Center, Houston; [4]Department of Medicine, Division of Rheumatology, University of California at San Francisco, San Francisco[†]; [5]Center for Pediatric Rheumatology, University of California at San Diego, La Jolla, CA[†]; [6]Program in Immunogenetics, Fred Hutchinson Cancer Research Center,[†] and [7]Department of Laboratory Medicine, University of Washington School of Medicine, Seattle; [8]Department of Medicine, University of Utah, Salt Lake City[†]; [9]Department of Medicine, Northwestern University Medical School, Chicago[†]; [10]Departments of Medicine and Microbiology, Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Birmingham, AL[†]; [11]Medical Research Service, Durham VA Hospital, Division of Rheumatology, Allergy, and Clinical Immunology, Duke University Medical Center, Durham, NC[†]; [12]Institute of Biological Anthropology, University of Oxford, Oxford; [13]Arthritis and Rheumatism Branch, National Institutes of Health/NIAMS,[†] and [14]CDR, MC, USN, Department of Radiology, National Naval Medical Center, Bethesda; [15]Department of Medicine, Vanderbilt University, Nashville[†]; and [16]Department of Orthopedics and Medicine, Thurston Arthritis Research Center, University of North Carolina, Chapel Hill, NC[†]

Rheumatoid arthritis (RA) is an autoimmune/inflammatory disorder with a complex genetic component. We report the first major genomewide screen of multiplex families with RA gathered in the United States. The North American Rheumatoid Arthritis Consortium, using well-defined clinical criteria, has collected 257 families containing 301 affected sibling pairs with RA. A genome screen for allele sharing was performed, using 379 microsatellite markers. A nonparametric analysis using SIBPAL confirmed linkage of the HLA locus to RA ($P < .00005$), with $\lambda_{HLA} = 1.79$. However, the analysis also revealed a number of non-HLA loci on chromosomes 1 (D1S235), 4 (D4S1647), 12 (D12S373), 16 (D16S403), and 17 (D17S1301), with evidence for linkage at a significance level of $P < .005$. Analysis of X-linked markers using the MLOD method from ASPEX also suggests linkage to the telomeric marker DXS6807. Stratifying the families into white or seropositive subgroups revealed some additional markers that showed improvement in significance over the full data set. Several of the regions that showed evidence for nominal significance ($P < .05$) in our data set had previously been implicated in RA (D16S516 and D17S1301) or in other diseases of an autoimmune nature, including systemic lupus erythematosus (D1S235), inflammatory bowel disease (D4S1647, D5S1462, and D16S516), multiple sclerosis (D12S1052), and ankylosing spondylitis (D16S516). Therefore, genes in the HLA complex play a major role in RA susceptibility, but several other regions also contribute significantly to overall genetic risk.

Introduction

Rheumatoid arthritis (RA [MIM 180300]) is a chronic, systemic, inflammatory disease with autoimmune features, the etiology of which remains unsolved. A genetic component to RA susceptibility has long been established by data from twin and family studies. A recent twin analysis estimated the heritability of RA to be ~60% (MacGregor et al. 2000). As in other complex autoimmune disorders with relatively low levels of familial aggregation, the strength of the genetic component has also been estimated by computing the value of λ_s, which is the relative recurrence risk for siblings of RA probands compared with that for the general population (Risch 1990*a*). In part because of difficulties with disease definition in population studies, the estimates of λ_s for RA vary widely, and range from 2 to 17 (Seldin et al. 1999). A consistent association between RA and human leukocyte antigen (HLA)–linked genes has been observed in many populations (Ollier and Thomson 1992). Although current evidence points to a role for HLA-

Received December 22, 2000; accepted for publication February 13, 2001; electronically published March 9, 2001.

Address for correspondence and reprints: Dr. Peter K. Gregersen, Division of Biology and Human Genetics, North Shore University Hospital, 350 Community Drive, Manhasset, NY 11030. E-mail: peterg@nshs.edu.

* All authors are members of the North American Rheumatoid Arthritis Consortium (NARAC).

[†] NARAC recruitment center.

© 2001 by The American Society of Human Genetics. All rights reserved.
0002-9297/2001/6804-0016$02.00

DRB1 alleles, the exact identity of the HLA susceptibility genes has not been established with certainty. It has been suggested that the HLA region genes account for a substantial fraction of the genetic component, perhaps as much as 50% (Deighton et al. 1989). Nevertheless, the balance of the evidence indicates that genes outside the HLA region also contribute to RA susceptibility.

In this study, we present the results of a genomewide screen and affected sib-pair analysis performed on 257 multicase RA families. These families were gathered by the North American Rheumatoid Arthritis Consortium (NARAC) for the first large multicenter study of RA multiplex families based in the United States (Gregersen 1998). This study was undertaken specifically to identify genes outside of the HLA region; given the increasing evidence for familial aggregation and genetic overlap among different autoimmune diseases (Becker et al. 1998), we also looked for evidence of common mapping results with other large linkage-mapping studies of autoimmune diseases.

Subjects and Methods

NARAC Family Collection

Multiplex RA families were recruited nationwide in the United States through the 12 participating recruitment centers of NARAC. Informed consent was obtained from every subject, including all participating family members, and the local institutional review board's approval was secured at every recruitment site prior to the start of enrollment. The criteria for entering a family into the study were as follows: (1) two or more siblings satisfied the 1987 American College of Rheumatology criteria for RA (Arnett et al. 1988), (2) at least one of the siblings had documented erosions on hand X-ray, and (3) at least one of the siblings had disease onset between the ages of 18 and 60 years. The presence of psoriasis [MIM 177900], inflammatory bowel disease (IBD [MIM 266600]) or systemic lupus erythematosus (SLE [MIM 152700]) was an exclusionary criterion in affected individuals. Radiographs of both hands were taken for all affected siblings at the time of recruitment into the study, unless films taken within 2 years were available. All the radiographs were scored by a single radiologist (D.F.), who used a preliminary, locally developed severity score using a scale of 0–5. The following criteria were used for this scale: 0, no erosion; 1, subtle erosion; 2, erosion without joint-space loss; 3, erosion with joint space loss; 4, complete joint-space loss, without marked bone destruction or joint deformity; and 5, marked bone destruction, with or without joint deformity. All patients were examined by trained NARAC personnel at time of entry and were scored for joint swelling, tenderness, alignment, and range of motion, according to standard published methods (Spiegel et al. 1987).

The recruitment of RA multiplex families by NARAC is ongoing, and we plan to repeat the analysis described in this study in independent data sets, to replicate the potential linkages we describe here. The summary data for all markers used in this genome screen can be found through links at the NARAC Web site. In addition, investigators may obtain access to individual-level genotyping data, digitized X-rays, and associated clinical data, after a brief application process. Limited quantities of DNA samples from these families are also available, upon request and approval, from the NARAC DNA repository, as described at the NARAC Web site.

Laboratory Procedures

Blood samples were collected, for DNA and serum, from the affected siblings and from their parents, if available. DNA was isolated from the peripheral mononuclear cells using a salting-out kit from BIO-101. At the Department of Laboratory Medicine at the University of Washington, rheumatoid factor (RF) was measured in each affected sibling by latex-enhanced nephelometric assay (Behring Diagnostics), which uses human and rabbit IgG coated on latex beads as antigen. This assay was calibrated to the World Health Organization's international standard for RF (Anderson et al. 1970). Broad-level typing for HLA-DRB1 was performed using allele-specific PCR to determine HLA allelic groups DRB1*01 through DRB1*18 (Olerup and Zetterquist 1998).

A genomewide screen was performed on 257 families (table 2), using markers, spaced at ~10-cM intervals, from the set8A combo list from the Marshfield map. Markers were added to the set at certain chromosomal locations, such as the HLA complex. The entire marker set consisted of 47 panels (379 markers), each containing markers pooled together according to size and fluorescent label (6-FAM, HEX, and NED). Reaction conditions were standard for all markers used, as described in the Marshfield PCR protocol (at the Marshfield Web site). Each panel of pooled markers was electrophoresed on a 310 or 3700 DNA Analyzer (Applied Biosystems).

Data Analysis

Semiautomated sizing of alleles was performed using the GENESCAN program, version 2.1 (Applied Biosystems), and individual genotypes were assigned with the help of the GENOTYPER program, version 2.1 (Applied Biosystems). To ensure accuracy of the genotypes, two individuals manually checked each genotype. We also ensured that this data set did not contain any MZ twin pairs, by searching the database for siblings with

Table 1

Structure of Families in the NARAC Collection

Parents Available	Number of Families			
	Overall (n = 257)	With Two Affected Sibs (n = 231)	With Three Affected Sibs (n = 24)	With Four Affected Sibs (n = 2)
Both	36	32	4	0
One	92	84	8	0
None	129	115[a]	12	2

[a] From one of these 115 families, five unaffected siblings were available for analysis. Otherwise, information from unaffected siblings was not used.

identical genotypes. Five such twin pairs were identified, and their dates of birth confirmed that they were twins. We also verified the genetic relationships among individuals by applying the program RELATIVE (Göring and Ott 1997) to all the marker data. In this manner, we identified nine half-sibling pairs and two unrelated pairs that were originally ascertained as full-sibling pairs.

A nonparametric linkage analysis was performed, using SIBPAL from the S.A.G.E. version 3.1 package for all autosomal marker data. We included all data from affected sib pairs to implement one-sided statistical t-tests of whether the identity-by-descent sharing departs from expectations when Mendelian inheritance of the markers is assumed. A further analysis of those chromosomes with regions showing some evidence of linkage ($P < .005$) was performed by examination of all markers on the relevant chromosomes using the ASM program of GENEHUNTER PLUS, under the Kong and Cox exponential model. We also analyzed the data allowing for the possible joint impact of HLA-region loci and other loci by weighting the Kong and Cox test statistics (Cox et al. 1999) using the family-specific NPL score at the chromosome 6 marker, D6S273. For markers on the X chromosome, the MLOD method provided by the SIBPHASE subroutine of ASPEX (Holmans 1998) was used. We computed a test statistic from the MLOD result by converting it to a χ^2 statistic which we assumed followed a 1:1 mixture of χ^2 deviates having 0 and 1 df, respectively. We also used the SIBIBD subroutine of ASPEX to calculate the probability (r_0) that sib pairs share 0 alleles identical by descent, given marker data for a location. The relative recurrence risk (λ_m) attributable to effects from a locus, m, that is tightly linked to genetic markers showing evidence for linkage, was calculated as $0.25/r_0$ (Risch 1990*b*). Because we used a rather sparse map with a marker interval of ~10 cM, these estimates are likely to be attenuated. With a 10-cM marker map, the maximum distance from the nearest marker to a disease-causing locus is 5 cM, and the average distance is 2.5 cM. To correct for a decrease in λ_m that can be expected from use of a sparse map, we calculated an unattenuated relative recurrence risk, λ_m'. The formula for λ_m' is based on formula 24 of Risch (Risch 1990*b*). The unattenuated relative recurrence risk is:

$$\lambda_{m'} = \left\{1 + \frac{1}{(1-2\theta)^2}\frac{1-\lambda_m}{\lambda_m}\right\}^{-1},$$

where θ is the recombination fraction, which we have set to .025. Descriptive statistics for demographic and clinical parameters were produced using the SAS package.

Results

RA Families

A total of 257 multicase families, containing 301 sibling pairs, were collected and analyzed for this initial genome screen. The distribution of affected siblings in these families is shown in table 1; in 49.8% of families, at least one parent was available for analysis. In addition to the affected siblings from whom DNA was obtained for the genomewide screen, there were other siblings in these families who did not participate in this study. Information on the number of siblings in each family nevertheless was available, as were clinical details if the siblings had been diagnosed with RA. We found that 7.7% of siblings (excluding the affected sib pair required for ascertainment of a family) were affected with RA. Under the assumption of a 1% prevalence of RA in the population, this recurrence risk yields a relative recurrence risk (λ_s) of 7.7. It should be noted, however, that this figure is a minimum estimate of the relative recurrence risk, since the estimates of the population prevalence vary and may be <1% (Seldin et al. 1999). On the other hand, the enrollment process for these families may be biased toward a higher rate of affected siblings. For example, families with multiple affected relatives may be more likely to include at least one sib pair willing to

Table 2

Clinical and Demographic Features of Affected Individuals (n = 546) in the NARAC Collection

Clinical/Demographic Feature	Value
Female	77.4%
White	90.0%
RF-positive	82.5%
HLA-DRB1*04 positive	70.0%
Hand erosions	94.0%
Mean age at disease onset (years)	39.0
Mean JAM score	29.9
Mean disease duration (years)	16.0

participate in the study, so that participation bias may inflate the recurrence-risk estimate.

The clinical features of the RA patients in these families are summarized in table 2. Approximately 82% of affected individuals had seropositive disease, and 70% carried at least one HLA-DRB1*04 allele. The mean age at onset of disease was 39 years, with mean disease duration of 16 years. The relatively long mean disease duration was reflected in clinical measures of cumulative joint damage. Each affected sibling was evaluated using a modified JAM score (Spiegel et al. 1987). The mean JAM score was 29.9 (range 0–112) of a maximum possible score of 120. As noted in the Subjects and Methods section, hand radiographs on all affected siblings were read and scored on a scale of 0–5 by a single collaborating radiologist. Using this method, the patients with erosions had a mean X-ray erosion score of 3.4.

Genomewide Screen

The results of the nonparametric analysis using SIBPAL are shown graphically in figure 1 and are summarized in table 3. Only markers with P values <.05 are included in table 3. As expected, markers in the HLA region (D6S265, D6S1629, and D6S273) showed significant sharing (0.56; $P < .00005$). A number of other regions on chromosomes 1 (D1S235), 4 (D4S1647), 12 (D12S373), 16 (D16S403), and 17 (D17S1301) showed linkage at a significance level of $P < .005$. On the X chromosome, there was marginally significant evidence for linkage to GATA52b03, with a P value of .016. In general, the analysis using the Kong and Cox exponential model showed evidence of linkage with the same chromosomal regions as did the analysis using SIBPAL (fig. 2). However, weighting the Kong and Cox LOD scores by the NPL test statistic from D6S273 did not lead to more-significant findings for any previously identified loci. Interestingly, on chromosome 17, weighting by D6S273 results eliminated evidence for linkage, possibly suggesting an inverse relationship in risk conferred by putative genetic factors on these two chromosomes (results not shown). Further results of these analyses can be found at the NARAC Web site.

In addition to analyzing the entire data set for linkage, we stratified the sibling pairs to examine whether linkage observed in the entire data set was contributed mainly by any one of the following subsets: (1) both siblings had disease onset before age 45 years, (2) the siblings were white, or (3) both siblings were seropositive for RF (RF >11 IU/ml). We selected the age of 45 years as a cut-off for early-onset disease, given that the mean age at onset for RA among first-degree relatives in multiplex families has been reported as 46 years (Lynn et al. 1995). This subgroup, where both affected siblings in a pair had disease onset before age 45 years, did not make any disproportionate contribution to the linkage demonstrated in the entire data set. Among the white families ($n = 230$), three markers showed a much higher level of significance than they did among all subjects (D3S3038, D3S2409, and D17S1308; $P < .005$ within the subgroup). The markers showing stronger evidence for linkage in the seropositive subgroup ($n = 185$) were D5S1462 and D12S373 ($P < .005$ within the subgroup). The latter marker exhibited considerable improvement in significance ($P = .00004$ within the subgroup) compared with the full data set ($P = .0031$ in the full set). To evaluate the significance of this finding, we performed some further analytical studies. We calculated the correlation among tests for each pair of subgroups and found the average correlations among the subgroup tests to be 0.6256. Then we found the critical value corresponding to $P = .00004$, which is 3.944. Finally, we obtained the nominal significance level corresponding to a three-dimensional test with this critical value, and the common correlation among the tests using software from Dunnett (STATLIB; Dunnett et al. 1989). By set-theory rules, this nominal significance is three times the nominal significances of the univariate tests, minus three times the bivariate tests, plus the trivariate test. The resulting P value is .00007. Overall, we had conducted 33 tests, and we had only tested regions that showed significance at the .05 level. For significance using these 33 tests and Sidak's (1967) correction, $P = 1 - (1 - p)^n$, we required a P value <.0000759; that is, $0.0025 = 1 - (1 - 0.000759)^{33}$. Technically, the subgroup finding for D12S373 is just significant, at the .05 level, after correction for multiple tests and for the fact that we required a significant result before further study. However, follow-up studies are needed to confirm this finding, both to assure a biologically significant finding and because significance tests are not very reliable in evaluation of extremely small P values.

Discussion

We report here the first major genomewide screen on multiplex families with RA that were gathered in North America. We used 379 markers at intervals of ~10 cM in 257 families with erosive RA. One important feature of the families recruited by NARAC is that all affected siblings were examined by trained NARAC personnel using a set of well-defined protocols, thereby mitigating uncertainty and potential lack of uniformity in clinical data collection. Moreover, erosion on the hand X-rays of at least one of the affected siblings in a family was a criterion for entry into the study; this was done in order to enhance confidence in the diagnosis of RA, as well as to provide a measure of disease severity. Overall, 82% of patients were seropositive for RF, and 70% carried at least one HLA-DRB1*04 allele. Our patients also

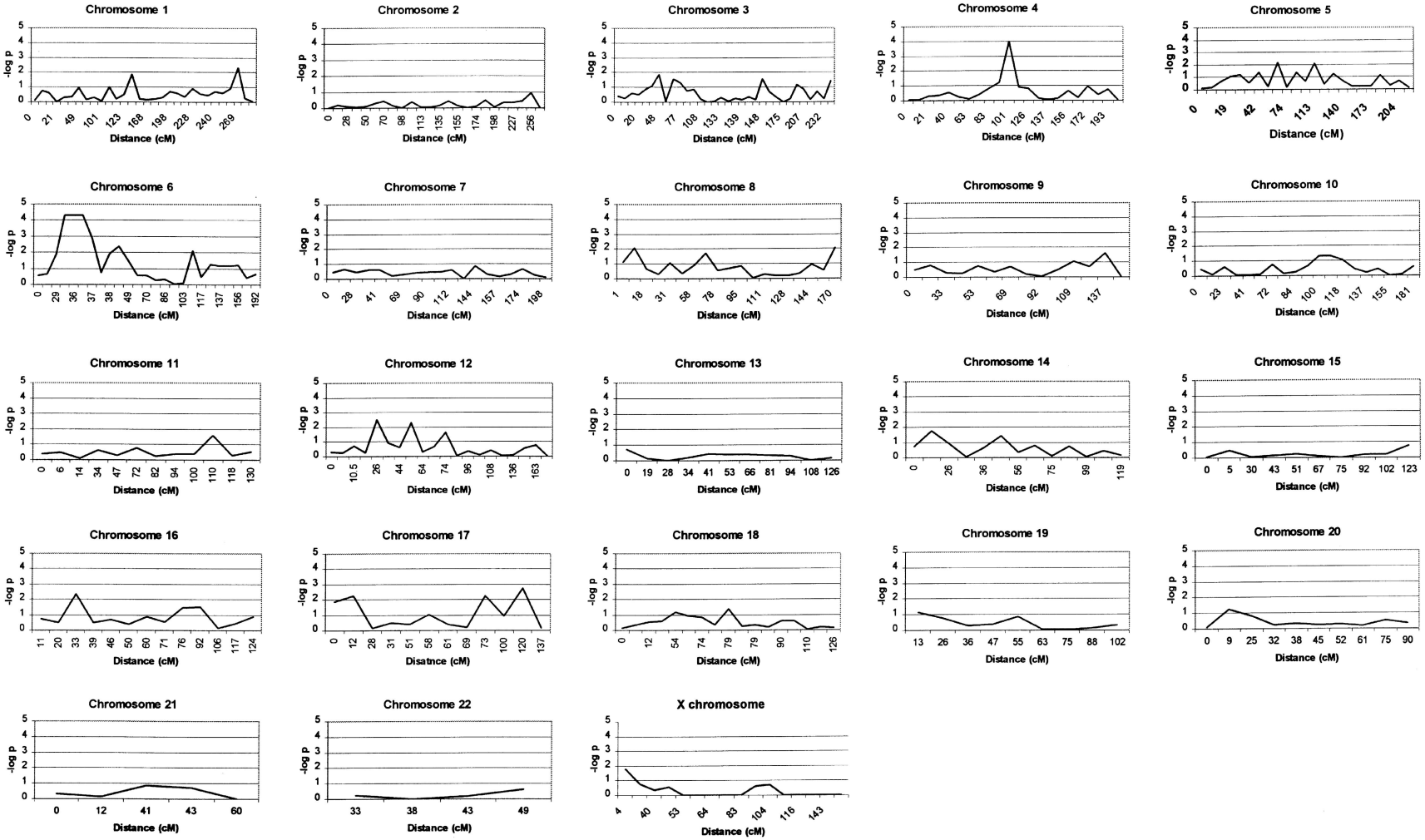

Figure 1 Results from affected sib-pair tests implemented by SIBPAL. The negative logarithm of the *P* values are plotted against the markers by chromosomal position. For markers on the X chromosome, the analysis was performed using the MLOD method from ASPEX; for this chromosome, too, the negative logarithm of the *P* values are plotted against the chromosomal position of the markers.

Table 3

Chromosomal Regions Giving *P* Values <.05 When SIBPAL 3.1 Is Used

Locus	No. of Sib Pairs	Mean Sharing	*P*	Distance[a] (cM)
Chromosome 1:				
D1S1631	237	.5426	.0141	137
D1S235	291	.5465	.0048	269
Chromosome 3:				
D3S3038	279	.5367	.0139	48
D3S1768	289	.5312	.0283	66
D3S2409	296	.5271	.0491	77
D3S4523	294	.5305	.0288	148
D3S1311	192	.5361	.0408	241
Chromosome 4:				
D4S1647	295	.5622	.0001	101
Chromosome 5:				
D5S2848	291	.5259	.043	42
D5S1457	290	.5395	.0066	63
D5S1501	253	.5277	.0411	92
D5S1462	286	.5378	.0079	113
Chromosome 6:				
D6S1959	267	.5336	.0135	29
D6S265	298	.574	.0000024	34[b]
D6S1629	299	.5604	.0000297	36[b]
D6S273	297	.5706	.000003	37[b]
D6S291	302	.5464	.0013	37[b]
D6S2439	284	.5472	.0043	40
D6S2427	291	.5315	.031	49
D6S1021	291	.5346	.0075	111
Chromosome 8:				
D8S277	293	.536	.0088	4
D8S1110	289	.5357	.0187	70
D8S373	285	.539	.0084	170
Chromosome 9:				
ATA59H06	252	.5368	.0244	137
Chromosome 10:				
D10S2327	267	.5276	.0462	100
D10S2470	286	.5276	.0476	113
Chromosome 11:				
D11S1998	263	.5271	.0258	110
Chromosome 12:				
D12S373	296	.5394	.0031	26
D12S398	297	.5393	.0051	57
D12S1052	288	.532	.0227	74
Chromosome 14:				
D14S1280	280	.5298	.0174	11
D14S587	267	.5305	.0373	46
Chromosome 16:				
D16S403	287	.5473	.0042	33
D16S2624	294	.5278	.0339	76
D16S516	300	.53	.0316	92
Chromosome 17:				
D17S1308	287	.529	.0137	0
D17S1298	285	.5348	.0053	12
D17S2180	289	.5363	.0054	73
D17S1301	293	.5431	.0017	120
Chromosome 18:				
D18S858	295	.5266	.0433	79

NOTE.—The means test was applied to the affected sib pairs to evaluate the one-sided hypothesis that identity-by-descent sharing exceeds the Mendelian expectation of 50%.

[a] Distances are from the Marshfield map, except where indicated.

[b] Distance from the Généthon map.

have a relatively young mean age at onset, 39 years, possibly due in part to the requirement that at least one of the affected siblings had disease onset between the ages of 18 and 60 years. Thus, overall, the NARAC collection of RA families contains affected siblings with relatively early onset and moderately severe erosive disease.

In order to confirm that nonparametric sib-pair analysis in our study population would allow us to detect linkage, we analyzed a few markers from the HLA region, since this region has previously been linked to RA. The results for the HLA markers confirmed this linkage with RA ($P < .00005$); as shown in table 3, markers D6S265 and D6S273 achieved significant linkage at the level suggested by Lander and Kruglyak ($P < 2.2 \times 10^{-5}$) (Lander and Kruglyak 1995). The relative contribution of HLA (λ_{HLA}) to the overall sibling recurrence risk for RA (λ_s) was found to be 1.79 in this data set. This figure for λ_{HLA} agrees with that found in the RA families recruited by the European Consortium for Rheumatoid Arthritis Families (ECRAF) (Cornélis et al. 1998).

Using the entire data set, we also calculated the marker-specific relative recurrence risks for the non-HLA region loci that showed evidence for linkage ($P < .005$), and these were as follows: $\lambda_{D1S235} = 1.14$, $\lambda_{D12S373} = 1.32$, $\lambda_{D16S403} = 1.25$, $\lambda_{D17S1301} = 1.32$, and $\lambda_{D4S1647} = 1.32$. After correction for attenuation, the relative recurrence risks were $\lambda_{D1S235} = 1.16$, $\lambda_{D12S373} = 1.37$, $\lambda_{D16S403} = 1.28$, $\lambda_{D17S1301} = 1.37$, and $\lambda_{D4S1647} = 1.37$. Because we used dense markers in the HLA region, we used a value of 1.79 for λ_{HLA} for further calculations using either attenuated or unattenuated recurrence risks. Risch (1990*b*) suggested that if multiple loci interact multiplicatively to affect disease risk, the sibling relative recurrence risk is determined by the product of the marker-specific relative recurrence risks. The manner in which different loci interact to increase disease risk for RA is unknown; however, the relative rarity of extended families with RA suggests that a multiplicative interaction is more consistent with the disease process than is an additive interaction. The products of the attenuated and unattenuated relative recurrence risks for all loci are 5.87 and 6.83, respectively. These relative recurrence risks are similar to the observed sibling recurrence risk of 7.7 in these families. Although the HLA region has the largest single impact on the relative recurrence risk, the joint effects of non-HLA region genes appear to be greater than those of the HLA region alone.

Lander and Kruglyak have suggested that significance in such genomewide screens should be at the level of $P < 2.2 \times 10^{-5}$ for significant linkage. This criterion corresponds to a 5% genomewide significance for a dense map and completely informative markers (Lander

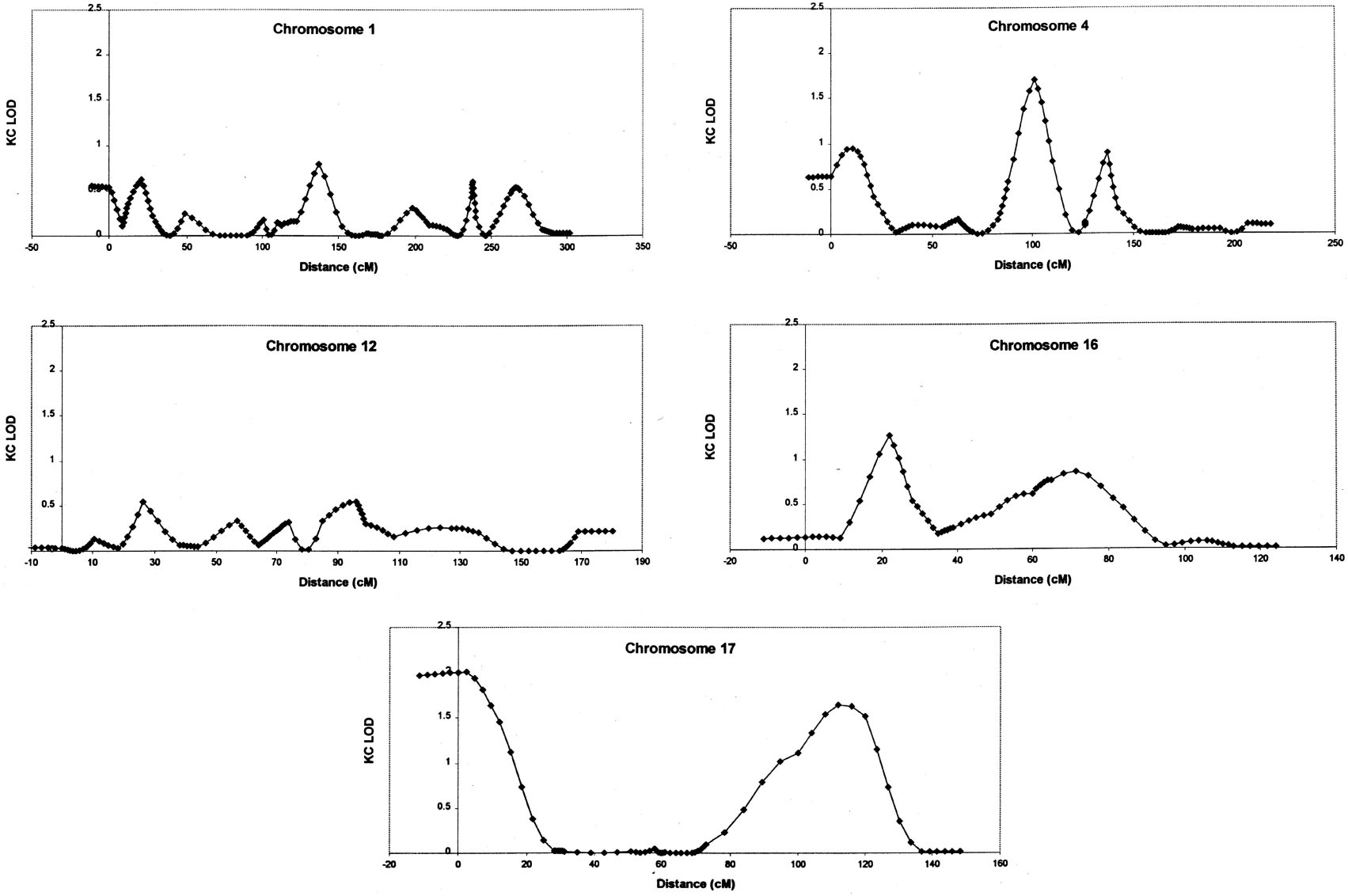

Figure 2 Results of the ASM analysis (GeneHunterPlus), under the Kong and Cox exponential model. Only the five chromosomes (1, 4, 12, 16, and 17) that show evidence of linkage in SIBPAL ($P < .005$) were considered for this analysis. The LOD scores are plotted against the chromosomal positions of the markers.

and Kruglyak 1995), which are conditions not met by our study. Overall, excluding the 80-cM region flanking the HLA complex, we conducted 350 tests using either SIBPAL or MLOD to examine whether the mean proportion of alleles shared identical by descent in affected siblings differed from 0.5. We identified 33 tests that were significant at the .05 nominal level. If we conservatively assume that each test is independent, we would expect to find ~18 linkages to be significant at the .05 level, and so we identified more putative linkages than expected if all linkages were false-positive findings. If we conservatively assume independence of all tests and then apply the Sidak (1967) correction for multiple tests, $P = 1 - (1 - p)^n$, at the genomewide significance levels of 5% and 10%, the critical values that we require are 1.5×10^{-4} and 3.0×10^{-4}. By these criteria, markers in the HLA region and D4S1647 are significant, whereas D17S1301 reaches a suggestive level of significance. To assess how correlated tests were with the nearby markers, we regressed the test statistics from each marker on the test statistics for the two adjacent markers to either site of a marker. The result of this analysis showed that ~15% of the variability among tests can be explained by the tests from the four flanking markers for any particular marker. Because the results from multiple test adjustments are not markedly affected by low levels of correlation among the tests (Brown and Russell 1996), application of a Bonferroni correction is likely to yield a reasonably conservative adjustment for multiple testing in our study. In a prior simulation (Zhao et al. 1999), it was found that the significance corresponding to the LOD score of 3.6 suggested by Lander and Kruglyak was too conservative, especially when a sparse map and many meioses were available for study. In reporting potential linkages, we have used a nonstringent cut-off of $P < .05$, to not miss signals from genes that are contributing only a modest effect. This was done to enable us to target potential linkages (Thomson 1994; Lander and Kruglyak 1995) that we will later confirm in independent data sets with even larger sample sizes. Using this nonstringent P value, we observed linkage on 14 different chromosomes, as summarized in table 3.

A stratification of the family data suggested that the white and seropositive subgroups may make disproportionate contributions to allele sharing at some loci. Among the white families, two markers on chromosomes 3 (D3S3038 and D3S2409) and one on chromosome 17 (D17S1308) showed an improved level of significance, compared with the entire data set. The families recruited by NARAC are ethnically heterogeneous, although 90% of the families used in this screen were white. The remaining subgroups were primarily Hispanic (5%), black (3.5%), Asian (0.8%), and American Indian (0.8%). The results in the white subgroup point to the possibility that there may be differences in RA-susceptibility loci between different ethnic groups. Indeed, this has been suggested by differences in patterns and strengths of HLA associations in different ethnic groups (Ollier and Thomson 1992). For example, black and Hispanic populations with RA do not exhibit a strong association with the HLA shared epitope (McDaniel et al. 1995; Teller et al. 1996). Thus, it is reasonable to assume that a degree of genetic heterogeneity is likely to exist at non-HLA susceptibility loci as well.

The results for the subgroup seropositive for RF (61% of affected sibling pairs) were the most striking. Marker D12S373 showed evidence for suggestive linkage at a level comparable to that of the HLA locus ($P < .00005$) in this subgroup. Given the evidence for the involvement of the tumor necrosis factor (*TNF*) pathways in the pathogenesis of RA, it is notable that this marker lies within 10 cM of the tumor necrosis factor receptor 1 (*TNFR1*) locus.

The linkage found at 16q24.1 in our data set has also been reported by ECRAF (Cornélis et al. 1998). That study is the only other large genomewide screen for RA-susceptibility genes that has been published. In addition, a recent meeting report on a candidate-region analysis of 152 RA sibling pairs indicates that a chromosomal region (17q22) containing marker D17S1301 also shows evidence of linkage to RA in families from the United Kingdom (Barton et al. 2000). It is of interest that this region is among the strongest linkages revealed by the Kong and Cox exponential model (fig. 2). We failed to replicate earlier reports from a Japanese genome screen showing linkage between markers on 1p36 and RA (Shiozawa et al. 1998). We also did not find significant evidence for linkage to the region of 3q13 identified by ECRAF (Cornélis et al. 1998).

RA is often grouped with a family of diseases classified as "autoimmune" that may share a common pathogenesis, which implies that a common group of genes may underlie autoimmunity (Becker et al. 1998). Therefore, we compared the markers that showed nominal significance in the present study with those reported as showing linkage to other autoimmune diseases. Five of the non-HLA chromosomal regions we report in this study have been implicated in at least one other autoimmune disease. The region of marker D1S235 ($P = .0048$) in the 1q43-44 region of chromosome 1 was first implicated in susceptibility to SLE by a candidate-region approach (Tsao et al. 1997). This region has been linked to SLE in three independent genome screens (Tsao et al. 1997; Gaffney et al. 1998; Moser et al. 1998). Marker D4S1647 on chromosome 4q, which demonstrated the most significant linkage with RA (outside the HLA region) in our data set ($P = .0001$), and marker D5S1462 ($P = .0079$) have shown evidence for linkage in families with IBD (Cho et al. 1998). Marker D12S1052 ($P =$

.0227) at 12q21.2 has been linked to multiple sclerosis (MS [MIM 126200]) (Haines et al. 1996). Another chromosomal region that raises interest is 16q24.1. Markers in that region have been implicated in IBD (Cho et al. 1998) and ankylosing spondylitis (AS [MIM 106300]; Brown et al. 1998), as well as in RA as described above (Cornélis et al. 1998).

As noted already, there are a number of regions of potential overlap between this study and other RA linkage studies (Cornélis et al. 1998; Barton et al. 2000). There are also discrepancies. Some of these may be due to type I error or may stem from differences in the ethnic compositions of the various family collections. It is likely that underlying genetic heterogeneity may explain differences among the studies. Epistatic interactions clearly influence expression of disease in murine models of autoimmunity (Morel et al. 2000). If RA in humans reflects epistatic interactions, then large data sets will be required to delineate all genetic factors; smaller studies may find some, but not all, of the causative genetic factors. Stratification of RA family data sets on the basis of genetic, as well as clinical, characteristics may help reveal these interactions. In this regard, we have done a preliminary stratification of our families by HLA-DR4. This has not shown significant differences from the full genome screen. Collaborative studies among the groups performing RA genome screens will allow further evaluation of our positive linkage findings on chromosomes 1, 4, 12, 16, and 17.

Acknowledgments

We would like to thank the following people for their contributions to the work described in this article: Aarti Damle, Susan Dowbak, Silvia Jaeger, and Dong Chen, for sample preparations and genotyping; Dr. Dorothy Guzowski, for synthesis of oligonucleotides; Nancy Giannola and Nina Kohn, for data management; Dakai Zhu and Wenfu Wang, for enabling us to perform error checking of the data through the Web; Jianfang Chen and Qingsong Liu, for the binning and error-handling programs used; Tracy Costello, for data processing design; and Sally Kaplan, Mary Noelle Holly, Clarine Cleveland, Peggy Rasmussen, Dana DePew, Carol Blalock, Karen Rodin, Linda Ingles, Diane Amox, Kay Randall, Donna McGregor, Marianna Crane, Pat Cummins, Phyllis Daum, and Jennifer Pearce, for outstanding field work in recruiting RA families. We are indebted to Clay Kasow and the staff at Empatheon, Inc. (formerly PatternRx, Inc.), for creating and maintaining the NARAC Web site to support this project. This project has been funded, in large part, by funds provided by the National Arthritis Foundation and federal funds from the National Institutes of Health, acting through the National Institute of Arthritis and Musculoskeletal and Skin Diseases and the National Institute of Allergy and Infectious Diseases, under grant RO1-AR44222 and contract NO1-AR-7-2232. These studies were performed, in part, in the General Clinical Research Center, Moffitt Hospital, University of California, San Francisco, with funds provided by the National Center for Research Resources, grant 5 M01 RR-00079, U.S. Public Health Service. The results in this article were obtained by using the program package SAGE, which is supported by a U.S. Public Health Service Resource Grant (1 P41 RR03655), from the National Center for Research Resources. We would like to extend our thanks to all the patients and their families for participating in the NARAC and making this study possible.

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

ASPEX, ftp://lahmed.stanford.edu/pub/aspex/index.html (for ASPEX software package)

Marshfield, http://www.marshmed.org/genetics/ (for markers and PCR protocol)

NARAC, http://www.naracdata.org/ (for summary marker data, information about obtaining DNA samples, and further results of genomewide screen)

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for RA [180300], SLE [152700], IBD [266600]), MS [126200], and AS [106300])

Statlib, http://www.statlab.uni-heidelberg.de/mirrors/statlib/apstat/.index.html (for software by Dunnet [reference 251])

References

Anderson SG, Bentzon MW, Houba V, Krag P (1970) International reference preparation of rheumatoid arthritis serum. Bull World Health Organ 42:311–318

Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, Healey LA, Kaplan SR, Liang MH, Luthra HS, Medsger TA Jr, Mitchell DM, Neustadt DH, Pinals RS, Schaller JG, Sharp JT, Wilder RL, Hunder GG (1988) The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 31:315–324

Barton A, Eyre S, Myerscough A, Silman A, Ollier W, Worthington J (2000) A novel RA susceptibility locus on chromosome 17 identified using syntenic mapping approaches. Arthritis Rheum Suppl 43:1225

Becker KG, Simon RM, Bailey-Wilson JE, Freidlin B, Biddison WE, McFarland HF, Trent M Jr (1998) Clustering of non-major histocompatibility complex susceptibility candidate loci in human autoimmune diseases. Proc Nat Acad Sci USA 95:9979–9984

Brown BW, Russell K (1997) Methods correcting for multiple testing: operating characteristics. Stat Med 16:2511–2558

Brown MA, Pile KD, Kennedy LG, Campbell D, Andrew L, March R, Shatford JL, Weeks DE, Calin A, Wordsworth, BP (1998) A genome-wide screen for susceptibility loci in ankylosing spondylitis. Arthritis Rheum 41:588–595

Cho JH, Nicolae DL, Gold LH, Fields CT, LaBuda MC, Rohal PM, Pickles MR, Qin L, Fu Y, Mann JS, Kirschner BS, Jabs EW, Weber J, Hanauer SB, Bayless TM, Brant SR (1998) Identification of novel susceptibility loci for inflammatory bowel disease on chromosomes 1p, 3q, and 4q: evidence for

epistasis between 1p and IBD1. Proc Nat Acad Sci USA 95: 7502–7507

Cornélis F, Faure S, Martinez M, Prud'Homme J, Fritz P, Dib C, Alves H, et al (1998) New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. Proc Nat Acad Sci USA 95:10746–10750

Cox NJ, Frigge M, Nicolae DL, Concannon P, Hanis CL, Bell GL, Kong A (1999) Loci on chromosomes 2 (NIDDM1) and 15 interact to increase susceptibility to diabetes in Mexican Americans. Nat Genet 21:213–215

Deighton CM, Walker DJ, Griffiths ID, Roberts DF (1989) The contribution of HLA to rheumatoid arthritis. Clin Genet 36:178–182

Dunnett CW (1989) Multivariate normal probability integrals with product correlation structure. Algorithm AS251. Appl Stat 38:564–579 [correction Appl Stat 42:709]

Gaffney PM, Kearns GM, Shark KB, Ortmann WA, Selby SA, Malmgren ML, Rohlf KE, Ockenden TC, Messner RP, King RA, Rich SS, Behrens TW (1998) A genome-wide search for susceptibility genes in human systemic lupus erythematosus sib-pair families. Proc Nat Acad Sci USA 95:14875–14879

Göring HH, Ott J (1997) Relationship estimation in affected sib pair analysis of late-onset diseases. Eur J Hum Genet 5: 69–77

Gregersen PK (1998) The North American Rheumatoid Arthritis Consortium: bringing genetic analysis to bear on disease susceptibility, severity and outcome. Arthritis Care Res 11:1–2

Haines JL, Ter-Minassian M, Bazyk A, Gusella JF, Kim DJ, Terwedow H, Pericak-Vance MA, et al (1996) A complete genomic screen for multiple sclerosis underscores a role for the major histocompatibility complex. Nat Genet 13: 469–471

Holmans P (1998) Affected sib-pair methods for detecting linkage to dichotomous traits: review of the methodology. Hum Biol 70:1025–1040

Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241–247

Lynn AH, Kwoh CK, Venglish CM, Aston CE, Chakravarti A (1995) Genetic epidemiology of rheumatoid arthritis. Am J Hum Genet 57:150–159

MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K, Silman AJ (2000) Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 43:30–37

McDaniel DO, Alarcon G, Pratt PW, Reveille JD (1995) Most African-American patients with rheumatoid arthritis do not have the rheumatoid antigenic determinant (epitope). Ann Intern Med 123:181–187

Morel L, Croker BP, Blenman KR, Mohan C, Huang G, Gilkeson G, Wakeland EK (2000) Genetic reconstitution of systemic lupus erythematosus immunopathology with polycongenic murine strains. Proc Natl Acad Sci USA 97: 6670–6675

Moser KL, Neas BR, Salmon JE, Yu H, Gray-Mcguire C, Asundi N, Bruner GR, Fox J, Kelly J, Henshall S, Bacino D, Dietz M, Hogue R, Koelsch G, Nightingale L, Shaver T, Abdou N, Albert DA, Carson C, Petri M, Treadwell EL, James JA, Harley JB (1998) Genome scan of human systemic lupus erythematosus: evidence for linkage on chromosome 1q in African-American pedigrees. Proc Natl Acad Sci USA 95:14869–14874

Olerup O, Zetterquist H (1993) HLA-DR typing by polymerase chain reaction amplification with sequence-specific primers (PCR-SSP). In: Hui KM, Bidwell JL (eds) Handbook of HLA typing techniques. CRC Press, London, pp 149–173

Ollier W, Thomson W (1992) Population genetics of rheumatoid arthritis. Rheum Dis Clin North Am 18:741–759

Risch N (1990*a*) Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 46:222–228

——— (1990*b*) Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 46:229–241

Seldin MF, Amos CI, Ward R, Gregersen PK (1999) The genetics revolution and the assault on rheumatoid arthritis. Arthritis Rheum 42:1071–1079

Shiozawa S, Hayashi S, Tsukamoto Y, Goko H, Kawasaki H, Wada T, Shimizu K, Yasuda N, Kamatani N, Tagasuki K, Tanaka Y, Shiozawa K, Imura S (1998) Identification of the gene loci that predispose to rheumatoid arthritis. Int Immunol 10:1891–1895

Sidak (1967) Z Rectangular confidence regions for the means of the multivariate normal distributions. J Am Stat Assoc 62:626–633

Spiegel T, Spiegel JS, Paulus H (1987) The joint alignment and motion scale: a simple measure of joint deformity in patients with rheumatoid arthritis. J Rheumatol 14:887–892

Teller K, Budhai L, Zhang M, Haramati N, Keiser HD, Davidson A (1996) HLA-DRB1and DQB typing of Hispanic American patients with rheumatoid arthritis: the "shared epitope" hypothesis may not apply. J Rheumatol 23: 1363–1368

Thomson G (1994) Identifying complex disease genes: progress and paradigms. Nat Genet 8:108–110

Tsao BP, Cantor RM, Kalunian KC, Chen CJ, Badsha H, Singh R, Wallace DJ, Kitridoou RC, Chen S, Shen N, et al (1997) Evidence for linkage of a candidate chromosome 1 region to human systemic lupus erythematosus. J Clin Invest 99:725–731

Zhao LP, Prentice R, Shen F, Hsu L (1999) On the assessment of statistical significance in disease-gene discovery. Am J Hum Genet 64:1739–1753